Title

Implicit Bias Predicts Liking of Ingroup Members Who Are Comfortable With Intergroup Interaction.

Journal

Personality & social psychology bulletin, 45(4)

ISSN

0146-1672

Authors

Jacoby-Senghor, Drew S
Sinclair, Stacey
Smith, Colin Tucker
et al.

Publication Date

2019-04-01

DOI

10.1177/0146167218793136

Peer reviewed

Article

Implicit Bias Predicts Liking of Ingroup Members Who Are Comfortable With Intergroup Interaction

Personality and Social Psychology Bulletin
2019, Vol. 45(4) 603–615

DOI: 10.1177/0146167218793136
journals.sagepub.com/home/pspb
SAGE

Drew S. Jacoby-Senghor[1], Stacey Sinclair[2], Colin Tucker Smith[3], and Jeanine L. M. Skorinko[4]

Abstract

We test a novel framework for how ingroup members are perceived during intergroup interaction. Across three experiments, we found that, above and beyond egalitarian attitudes and motivations, White observers' automatic responses to Blacks (i.e., their implicit anti-Black bias) shaped their affiliation toward ingroup targets who appeared comfortable engaging in interracial versus same-race interaction. White observers' implicit anti-Black bias negatively correlated with liking of White targets who were comfortable with Blacks (Experiments 1-3). The relationship between implicit bias and liking varied as a function of targets' nonverbal comfort in interracial interactions (Experiment 1). Specifically, implicit bias negatively correlated with liking of targets when targets' nonverbal behaviors revealed observers felt comfortable with interracial contact, irrespective of the nature of those behaviors (Experiment 2). Finally, the relationship between implicit bias and target liking was mediated by perceived similarity (Experiment 3). Theoretical implications for stigma-by-association, social network homogeneity, and extended contact are discussed.

Keywords

implicit prejudice, intergroup relations, interpersonal perception, stigma-by-association

Received September 19, 2017; revision accepted July 15, 2018

Research has shown that one's level of comfort with intergroup interaction unintentionally gives rise to corresponding nonverbal behaviors (Dasgupta, 2004; Dovidio, Kawakami, & Gaertner, 2002; Greenwald, Poehlman, Uhlmann, & Banaji, 2009). In turn, these nonverbal cues of comfort affect a range of outcomes, including interactants' perceptions of one another (Word, Zanna, & Cooper, 1974) and their desire to affiliate (Dovidio et al., 2002). The present research turns from studying the experience of individuals within such interactions to take the perspective of individuals *observing* them. Specifically, we examine the degree to which White observers' automatic responses to Blacks shape their liking of Whites viewed interacting with a Black stranger. This is an important question because intergroup interactions do not occur in a vacuum but instead are observed by myriad individuals, such as acquaintances, coworkers, and mere passersby. A full understanding of intergroup interaction therefore involves not only what occurs within such exchanges but also how these exchanges reverberate among others within the immediate social milieu.

The closest researchers have come to studying this topic is investigating the social costs incurred by people who are observed associating with stigmatized individuals (i.e., stigma by association; Goffman, 1963). Research on stigma by association focuses almost exclusively on whether members of nonstigmatized groups devalue ingroup peers who have close relationships with stigmatized others (e.g., family, friends, and roommates; Neuberg, Smith, Hoffman, & Russell, 1994; Pryor, Reeder, & Monroe, 2012; Rozin, Markwith, & Nemeroff, 1992; Sigelman, Howell, Cornell, Cutright, & Dewey, 1991). However, people's actual social networks often still do not include outgroup members (Wimmer & Lewis, 2010), and observing close relationships that span group boundaries remains infrequent (Dixon, Tredoux, Durrheim, Finchilescu, & Clack, 2008). Therefore, the kinds of intergroup contact that form the focus of the literature at present (i.e., close relationships) are less prominent in everyday life than those that are neglected (e.g., interactions between strangers). Given these issues, conducting

[1]University of California, Berkeley, USA
[2]Princeton University, NJ, USA
[3]University of Florida, Gainesville, USA
[4]Worcester Polytechnic Institute, MA, USA

Corresponding Author:
Drew S. Jacoby-Senghor, Haas School of Business, University of California, Berkeley, 2220 Piedmont Avenue, Berkeley, CA 94720, USA.
Email: dsenghor@gmail.com

research on the perceptions of ingroup members who are apparently comfortable or uncomfortable during passing intergroup interactions is crucial.

A small subset of stigma by association research has investigated how associative stigma operates when observing contact between two people who are strangers (e.g., Hebl & Mannix, 2003; Pryor et al., 2012). This research has shown that mere proximity to a stigmatized individual can be sufficient for associative devaluation to occur. Hebl and Mannix (2003), for example, showed that participants devalued a male target viewed in mere proximity to an obese woman, even without any information about a relationship between the two. Yet, explanations for the "mere proximity" phenomenon remain speculative, and those that have been experimentally tested focus on cognitive mechanisms to the exclusion of social processes that may be occurring (e.g., Walther, 2002).

We propose a novel hypothesis that integrates the implicit bias and similarity-attraction literatures: Observers' liking of ingroup targets is driven by the correspondence between observers' automatic response to the outgroup and ingroup targets' comfort with intergroup interaction. Specifically, we predict that Whites' implicit anti-Black bias will negatively correlate with their expected liking of fellow Whites whom they see comfortably interacting with a Black stranger. In other words, Whites with more positive automatic responses to Blacks, compared with those with more negative responses, should show greater liking of ingroup members who are comfortable with a Black person. In contrast, when observing White targets who are uncomfortable with Blacks, the relationship between implicit bias and liking may yield one of two outcomes: One possibility is that implicit anti-Black bias may *positively* correlate with liking of these uncomfortable targets. However, another possibility is that people's general dislike of individuals who are socially uncomfortable (Alden & Wallace, 1995; Jones & Carpenter, 1986) and who flaunt egalitarian norms (Dovidio & Gaertner, 2004) will cause the relationship between implicit bias and liking to weaken to nonsignificance.

Forming the foundation for this hypothesis is an extensive literature demonstrating that implicit bias shapes individuals' comfort with intergroup interaction above and beyond explicit racial attitudes, which instead correspond with more deliberative responses (Nosek & Smyth, 2007). Whites' implicit bias is associated with how comfortable they report being around Blacks (Towles-Schwen & Fazio, 2003), their physiological stress during such interactions (Blascovich, Mendes, Hunter, Lickel, & Kowai-Bell, 2001), and their mental energy afterward (Richeson & Shelton, 2003). People's self-reported "gut feelings" toward outgroup members also correspond with their actual level of implicit bias (Hahn, Judd, Hirsh, & Blair, 2014; Ranganath, Smith, & Nosek, 2008). In sum, people appear able to reflect on and recognize their automatic responses to outgroup members despite these responses being difficult to suppress or control.

When observing ingroup members engaged in intergroup interactions, people are also able to identify signals of comfort. For example, Whites display subtle nonverbal behaviors of comfort during interracial interaction that are difficult to control yet readily and accurately interpreted by White third-party observers (Dovidio et al., 2002; Richeson & Shelton, 2005). Therefore, people not only have a sense of how comfortable they are with intergroup interaction but also how comfortable others are with such encounters. Taken together, these findings suggest that Whites possess ample means to infer whether their level of comfort with Blacks corresponds with another's. Furthermore, because implicit bias uniquely predicts one's comfort with Blacks, it should also uniquely predict the degree to which one's comfort apparently corresponds with that of an ingroup member.

A robust literature on similarity demonstrates that the degree of one's correspondence with another in turn affects the degree to which one is drawn to them. Research has shown time and again that people generally like others who have similar behaviors, attitudes, backgrounds, and preferences (Berscheid, 1985), including those specific to race (Conley, Rabinowitz, & Hardin, 2010). Although past work also has shown that *dissimilarities* between people can make them *less* likely to affiliate (Chen & Kenrick, 2002; Singh & Ho, 2000), the present work does not differentiate between the effects of similarity and dissimilarity but rather focuses on whether correspondence in comfort with the outgroup predicts affiliation at all. Most relevant to this goal, recent research demonstrates that Whites infer how similar they are to White targets from the race of targets' friends and subsequently prefer those whose friendship choices indicate a similar experience of interracial contact (Jacoby-Senghor, Sinclair, & Smith, 2015). Together, this research suggests that people are attuned to how much their automatic responses to the outgroup correspond with ingroup members' apparent comfort with the outgroup and that the level of correspondence should drive affiliation.

Current Experiments

Across three experiments, we asked whether White participants' implicit anti-Black bias predicted the degree to which they expected to like a White person observed interacting with a Black stranger. In Experiment 1, we manipulated the nonverbal behaviors of White targets to determine whether their apparent comfort or discomfort with interracial interaction changed participants' responses toward them. This design also tested our hypothesis against the alternative stigma by association explanation that observers will dislike ingroup targets for merely being in proximity to outgroup members (e.g., Hebl & Mannix, 2003) irrespective of the ingroup targets' apparent comfort. In contrast, we expected implicit anti-Black bias to negatively correlate with liking of White targets observed comfortably interacting with a Black stranger but for this relationship to disappear or reverse when targets were uncomfortable.

In Experiment 2, we manipulated what participants thought ingroup targets felt during the interactions to directly test whether perceptions of targets' comfort with interracial interaction drive the effects. Specifically, orthogonal to the valence of the nonverbal behaviors of targets, we told participants the behaviors they were observing indicated comfort or discomfort in the interaction. This design allowed us to separate whether participants' ratings were driven by the apparent correspondence between targets' comfort and participants' own automatic responses to the outgroup or whether ratings were instead driven by targets' biased behavior. As such, this experiment provides a contrast between our hypothesis and the alternative interpretation that more biased participants simply prefer ingroup members who behave prejudicially toward the outgroup (e.g., Castelli, De Amicis, & Sherman, 2007).

In Experiment 3, we tested our assumption that implicit bias will predict liking to the extent that Whites infer correspondence between the Whites' targets they are rating and themselves. To do so, we measured perceived similarity to ingroup targets to test whether it mediated the relationship between their implicit anti-Black bias and liking of targets.

Across the three studies, we tested whether participants' implicit anti-Black bias independently predicted expected liking above and beyond alternate forms of racial bias (i.e., explicit racial bias, egalitarian motivation, and implicit ingroup favoritism). In accordance with the precedent of the vast majority of research on implicit bias, we predict relative differences between those high and low in implicit bias and do not specify that either group drives the observed effects.

Experiment 1

In this experiment, we tested whether White participants' implicit anti-Black bias predicted the degree to which they expected to like a White person observed interacting with a Black stranger, as well as whether the nature of this relationship depended on how comfortable the White person being observed seemed. We expected implicit anti-Black bias to negatively correlate with liking of White targets observed comfortably interacting with a Black stranger. For uncomfortable targets, we identified two possible outcomes. On one hand, the proposed relationship might reverse with implicit bias and liking becoming positively correlated. On the other hand, these targets might be perceived as undesirable to such a degree that the effect weakens. Not only are uncomfortable individuals generally less liked (Alden & Wallace, 1995; Jones & Carpenter, 1986), but obvious expressions of prejudice are often shunned in light of present-day egalitarian norms (Dovidio & Gaertner, 2004), especially with respect to transgressive racial attitudes (Crandall, Eshleman, & O'Brien, 2002). Therefore, the relationship between implicit bias and liking might become nonsignificant when rating uncomfortable targets.

Confirmation of either of these two predictions for uncomfortable targets is notable because extant research has often argued that targets' nonverbal comfort with the outgroup may not matter. Previous stigma by association research, for example, contends that perceivers may dislike ingroup targets who are merely proximal to stigmatized outgroup members (e.g., Hebl & Mannix, 2003). Such a perspective predicts that more implicitly biased participants will dislike ingroup targets observed interacting with Blacks regardless of how comfortable or uncomfortable the targets appear. As such, finding that the apparent comfort of the target changes the relationship between implicit bias and liking would help rule out a stigma by association interpretation of the present work.

In addition, we expected the predicted effects to hold even when controlling for explicit egalitarian attitudes and motivations. We therefore measured explicit prejudice (Katz & Hass, 1988) and internal motivation to respond without prejudice (Plant & Devine, 1998).

Method[1]

Participants. We recruited White U.S. participants via Amazon Mechanical Turk. They were paid US$1.00 each. Given that interpersonal evaluations can differ as a function of whether the target is within or outside of one's own generational cohort (North & Fiske, 2013), we recruited participants who were within the range of the undergraduate targets' perceived age (i.e., 18 to 30 years old). As this was our first use of this paradigm on Mechanical Turk's online platform, we calculated our target sample size using an estimated effect size, *f*, of implicit bias on liking of 0.25, which would require a sample size of approximately 165 participants for the study to be powered at 90%. We therefore collected 169 participants (M_{Age} = 24.90 years, SD_{Age} = 3.24; 49% male). Fourteen participants were removed for having an average response latency of less than 1 s on their target ratings. Thus, 155 participants remained in the final sample (M_{Age} = 24.84 years, SD_{Age} = 3.17; 52% male).

Procedure. Participants were informed that they would watch eight silent videos of interactions between people meeting for the first time and received the following instructions: "For this task, we are interested in how friendly people come across as being. We have collected several video clips of two unacquainted people having a conversation for the first time. You will evaluate one person in each pair." All participants were told that they would rate the person on the right side of the screen. This target person was always White. After each video, participants rated their liking of the specified target. After rating all eight videos, they then completed the implicit and explicit measures and demographic questions.

Materials

Interaction videos. We selected a set of 16 videos converted from a past study on interracial interactions (Mallet, Wilson, & Gilbert, 2008) and shortened them to 20 s each.

Each participant saw a total of eight silent clips of undergraduate White female targets interacting with an undergraduate female partner. Videos were divided equally by partner race (Black vs. White) and target nonverbal comfort (high vs. low). Each participant saw eight videos, two from each of the four conditions. Four coders (one male) rated targets' nonverbal comfort on three dimensions (i.e., comfort, friendliness, and rapport; 1 = *not at all comfortable* to 7 = *very comfortable*; $\alpha = .86$). Ratings for each condition were as follows: comfortable: $M_{Black} = 5.6$, $SD_{Black} = 0.2$; $M_{White} = 5.4$, $SD_{White} = 0.4$; uncomfortable: $M_{Black} = 2.9$, $SD_{Black} = 0.7$; $M_{White} = 3.3$, $SD_{White} = 0.4$.

Our hypothesis states that participants infer the extent to which targets and they share a comparable level of apparent comfort with the outgroup. We therefore collected data on whether the race and apparent comfort of targets' interaction partners indeed affect participants' perceptions of targets' comfort with the outgroup. White participants ($N = 49$; one participant dropped for not completing the task) were recruited via Amazon Mechanical Turk and asked to rate the targets in each of the videos on three outgroup comfort items using a 6-point Likert-type scale ("This person likely feels uncomfortable around people from different backgrounds." reverse-scored; "This person likely has many friends from different backgrounds."; "This person likely feels nervous talking with people who are different from them." reverse-scored; 1 = *strongly disagree*; 6 = *strongly agree*; $\alpha = .82$). The items were averaged such that higher scores indicated higher perceived comfort with outgroup members. As predicted, paired *t* tests confirmed that targets who exhibited positive (i.e., comfortable) nonverbal behavior with a Black ($M = 4.82$, $SD = 0.91$), compared with White ($M = 4.49$, $SD = 0.88$), partner were seen as more comfortable with outgroup members in general, $t(48) = 2.82$, $p = .007$, 95% confidence interval [CI] = [0.10, 0.57]. Targets who exhibited negative nonverbal behavior (i.e., *un*comfortable) with a Black ($M = 3.06$, $SD = 0.87$), compared to White ($M = 3.38$, $SD = 0.72$), partner were seen as marginally less comfortable with outgroup members in general, $t(48) = -1.94$, $p = .059$, 95% CI = [−0.65, 0.01]. Therefore, our comfort manipulation had the intended effect on participants' perceptions of targets' comfort with outgroup members.

Liking of target. To assess liking, participants rated three items on a 6-point Likert-type scale (1 = *strongly disagree*/*not at all*; 6 = *strongly agree*/*very much*). These were "I would expect to get along easily with this person," "I would expect to have a smooth interaction with this person," and "To what extent do you think you would want to become friends with this person?" The three items demonstrated acceptable reliability ($\alpha = .91$).

Single Target Implicit Association Test (ST-IAT). AST-IAT(Bluemke & Friese, 2008) served as the measure of implicit anti-Black bias. Within the task, participants saw a single discrimination block of the evaluative stimuli (e.g., "marvelous"/"superb"/"pleasure" and "tragic"/"horrible"/"agony"). Next, participants completed a block in which monochromatic photos of Black faces and positive stimuli were indicated using one keyboard key and negative stimuli were indicated using another. The stimuli pairing was then reversed in the final block. The ST-IAT provided a score of evaluative associations for Blacks alone rather than a relative score as in the traditional IAT. Higher scores indicated greater implicit anti-Black bias.

Explicit racial bias measures. We used the 10-item Anti-Black subscale of the Ambivalent Racism scale (Katz & Hass, 1988; $\alpha = .88$) and measured participants' internal motivation to be non-prejudiced using Plant and Devine's (1998) five-item scale (i.e., Internal motivation to be non-prejudiced scale (IMS); $\alpha = .84$).

Results

Across all experiments, we analyzed whether participants' levels of implicit anti-Black bias predicted their reactions to Whites with a Black interaction partner relative to their ratings of Whites seen interacting with Whites. To test our hypothesis that implicit outgroup bias should predict ingroup affiliation above and beyond alternate forms of racial bias, analyses for all studies controlled for the effects of all alternative forms of racial bias that were measured.[2] Analyses were conducted with generalized estimating equations (GEE), specifying a robust estimator and exchangeable covariance structure. Appropriate for our repeated measures data and interest in differences between respondents with varying levels of implicit bias, GEE focuses on between-subject effects, rather than primarily on within-subject effects (Zeger & Liang, 1986). GEE analyses are robust to misspecifications of the correlation matrix and do not require any specific distribution among responses. Although GEE uses a Wald chi-square test statistic, coefficients from GEE models have analogous meaning to coefficients from standard multiple regression (e.g., Seery, Holman, & Silver, 2010). Effect size metrics in GEE are not well established, so rough effect size estimates were calculated for power analyses using more traditional, but less fitting, regression analyses.

Mean-centered implicit anti-Black bias, explicit anti-Black bias, and IMS scores were entered as predictors in the analysis. Comfort (comfortable = 0; uncomfortable = 1) and all two-way interactions between comfort and each type of bias were also entered as predictors. Higher scores indicated greater liking of targets interacting with Blacks relative to targets interacting with Whites. This approach allowed us to directly test whether different levels of implicit bias affect participants' likelihood of affiliating with targets viewed in interracial interaction compared with their baseline responses toward targets in same-race interaction. This strategy also follows recommended approaches when analyzing repeated-measures data with continuous predictors and categorical

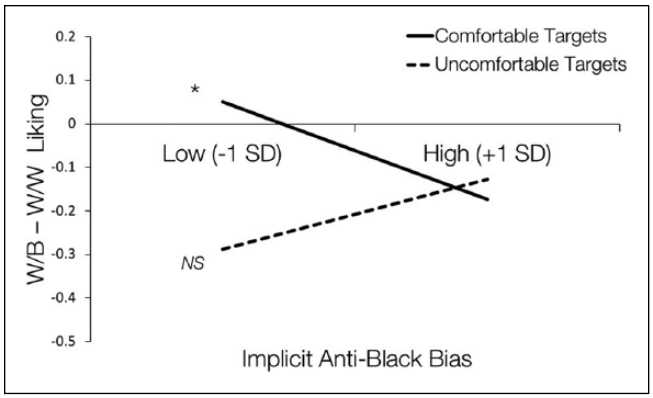

Figure 1. Experiment 1: Effect of implicit anti-Black bias on liking of ingroup targets nonverbally comfortable or uncomfortable with a Black (W/B) versus a White (W/W) interaction partner.
Note. Higher numbers indicate increasing preference for Whites with Blacks relative to Whites with Whites.

moderators (Judd, Kenny, & McClelland, 2001), and follows previous relevant stigma by association research (e.g., Pryor et al., 2012). Tables for all experiments, including descriptive data and statistics for nonrelative dependent variables, are presented in supplemental materials.[3]

The predicted two-way interaction of implicit anti-Black bias and comfort was significant ($B = .48$, $SE = .23$, Wald $\chi^2 = 4.31$, $p = .038$, 95% CI = [0.94, 4.31]). When participants observed comfortable interactions, their degree of implicit bias was negatively correlated with their liking of White targets interacting with a Black person ($B = -.28$, $SE = .12$, Wald $\chi^2 = 5.56$, $p = .018$, 95% CI = [−0.51, −0.05]; see Figure 1). However, when the interactions were uncomfortable, implicit bias did not predict liking of targets ($p = .274$).

Discussion

As hypothesized, Whites' implicit anti-Black bias predicted their liking of ingroup targets seen interacting with Blacks, and the nature of this relationship depended on how comfortable ingroup targets seemed. When participants observed fellow Whites comfortably interacting with a Black person, their degree of implicit bias was negatively correlated with liking of the White target. This effect not only persisted above and beyond participants' level of explicit prejudice but also when controlling for their motivation to control their prejudicial responses, suggesting they exerted limited control over expression of this preference.

As expected, this relationship differed when targets appeared uncomfortable with their interaction partner. We laid out two possibilities for what may occur when targets appeared uncomfortable: (a) Implicit bias and liking of White targets might be positively correlated, or (b) there would be no relationship between implicit bias and liking. We found the latter outcome, though the nonsignificant relationship between implicit bias and liking was positive. As previously discussed, this lack of relationship could have occurred for multiple reasons. For one, people perceived as uncomfortable during interactions are generally less liked (Alden & Wallace, 1995; Jones & Carpenter, 1986), and participants' general dislike for these targets may have overridden the effect. Indeed, participants liked uncomfortable targets ($M_{\text{Uncomfortable}} = 3.37$) much less than comfortable targets ($M_{\text{Comfortable}} = 4.46$); $t(154) = 16.66$, $p < .001$, 95% CI = [0.96, 1.22]. A second possible explanation is that indicating preference for targets who are apparently uncomfortable with outgroup contact may be aversive to participants. People often prefer to be seen as egalitarian when such attitudes are normative (Dovidio & Gaertner, 2004), and tend to monitor their responses when they believe they may appear prejudiced (Crandall & Eshleman, 2003; Vorauer & Turpie, 2004). Endorsing affiliation with potentially prejudiced targets may have heightened such concerns for participants to a greater degree than when rating targets apparently comfortable with outgroup contact.

Irrespective of the reason for the lack of relationship between implicit anti-Black bias and liking toward uncomfortable ingroup targets, finding that the relationship is moderated by targets' nonverbal behavior was in line with our hypothesis. Namely, observers do not judge the likability of ingroup members based only on their mere proximity to outgroup members; instead, the degree to which individuals like an ingroup target observed in interracial interaction depends on both the target's apparent comfort with such interactions and the observer's own automatic responses to the outgroup (i.e., their implicit anti-Black bias).

Experiment 2

We next directly tested the central hypothesis that liking of ingroup members is driven not simply by how targets act during intergroup contact but instead specifically by how comfortable targets are thought to feel. A plausible alternative explanation of Experiment 1 is that participants observing these interactions are not responding to the apparent comfort of ingroup targets but to the tenor of their behaviors per se. In other words, it could be the case that Whites' implicit bias is negatively correlated with liking of Whites who behave warmly toward Blacks (e.g., Castelli et al., 2007), as opposed to those who are comfortable around Blacks.

To test our hypothesis and against this alternative explanation, we orthogonally manipulated participants' perceptions of how comfortable targets were and the valence of targets' nonverbal behavior. Some participants were told that ingroup targets who feel comfortable do indeed appear nonverbally comfortable while other participants were told that ingroup targets who feel comfortable ironically appear nonverbally uncomfortable. We predicted that participants' implicit anti-Black bias would predict expected liking when targets were thought to *actually feel* comfortable but not

when their behaviors appeared comfortable but belied an uncomfortable internal state.

Method

Participants. We calculated our target sample size using an effect size, *f*, of implicit bias on liking of 0.38, which was the estimated average effect size across the previously run studies. Analysis suggested a sample size of approximately 90 participants for the study to be powered at 90%. We collected 89 White U.S. participants ranging from 18 to 30 years in age via Amazon Mechanical Turk (age $M = 24.95$ years, $SD = 3.08$; 61% male). Five participants were removed for having an average response latency of less than 1 s on their target ratings. Eighty-four participants remained in the final sample (age $M = 24.90$ years, $SD = 3.15$; 60% male). Participants were paid US$1.

Procedure. The procedure was the same as Experiment 1 with one exception: the instructions given to participants. After explaining that participants would view eight silent videos and evaluate one person in each video, participants were given additional instructions that led them to believe that the targets' behaviors matched their felt comfort or did not. In the match condition, participants learned, "Research shows that in initial interactions people's behavior usually corresponds to their inner states. For example, people who feel the most comfortable during a first-time interaction usually appear friendly. But, people who feel the most uncomfortable usually appear unfriendly." In the opposite meaning condition, participants instead learned, "Research shows that in initial interactions, people's behavior usually does not correspond to their inner states. For example, people who feel the most comfortable during a first-time interaction usually appear unfriendly. But, people who feel the most uncomfortable usually appear friendly."

The rest of the experiment proceeded identically to Experiment 1. After the manipulation, participants viewed the same set of videos from Experiment 1, rated expected liking for all targets ($\alpha = .91$), and completed the ST-IAT (Bluemke & Friese, 2008) as the measure of implicit anti-Black bias and the Anti-Black Bias subscale of the Ambivalent Racism scale ($\alpha = .89$) and motivation to respond without prejudice (i.e., IMS; $\alpha = .87$) as measures of explicit prejudice.

Results

Mean-centered implicit anti-Black bias, explicit anti-Black bias, and IMS scores were entered as predictors in the analysis. Nonverbal comfort (comfortable = 0; uncomfortable = 1), instructions condition (match condition = 0; opposite meaning condition = 1), and all two- and three-way interactions between nonverbal comfort, instructions condition, and each type of racial bias were also entered as predictors. Liking ratings for White targets with a Black interaction partner minus White targets with a White interaction partner was again entered as the dependent variable.[4]

The predicted three-way interaction of implicit anti-Black bias, nonverbal comfort, and instructions was significant ($B = -3.08$, $SE = 1.11$, Wald $\chi^2 = 7.71$, $p = .005$, 95% CI = [−5.26, −0.91]).

Match condition. As expected, when targets' nonverbal behavior and felt comfort were said to match, the results replicated those of the previous experiment. The predicted two-way interaction of implicit anti-Black bias and nonverbal comfort approached significance ($B = 1.88$, $SE = .97$, Wald $\chi^2 = 3.74$, $p = .053$, 95% CI = [−0.02, 3.79]). When participants observed comfortable interactions, their degree of implicit bias was negatively correlated with their liking of White targets interacting with a Black person ($B = -1.21$, $SE = .44$, Wald $\chi^2 = 7.60$, $p = .006$, 95% CI = [−2.07, −0.35]; see Figure 2a). However, when the interactions were uncomfortable, implicit bias did not predict liking of targets ($p = .354$).

Opposite meaning condition. When targets' nonverbal behavior and felt comfort were said to be in opposition, the predicted two-way interaction of implicit anti-Black bias and nonverbal comfort was also significant ($B = -1.20$, $SE = .53$, Wald $\chi^2 = 5.06$, $p = .024$, 95% CI = [−2.24, −0.15]). We next tested the prediction that implicit bias would predict ratings of targets who were said to feel comfortable even though they displayed negative nonverbal behaviors. As predicted, participants' implicit bias was negatively correlated with their liking of those White targets ($B = -1.01$, $SE = .45$, Wald $\chi^2 = 5.03$, $p = .025$, 95% CI = [−1.89, −0.13]; see Figure 2b). Finally, we examined ratings of targets who were said to feel *un*comfortable but displayed positive nonverbal behaviors. As expected, based on the null findings, when targets appeared uncomfortable in Experiment 1 and in the match condition, implicit bias did not predict liking for such targets ($p > .688$).

Discussion

In Experiment 2, we orthogonally manipulated participants' perception of how comfortable targets were and the valence of targets' nonverbal behavior. In doing so, we were able to provide additional evidence that participants' liking of ingroup targets is driven by the degree to which their own automatic responses to the outgroup correspond with targets' apparent comfort with outgroup contact. Replicating Experiment 1, when participants observed fellow Whites they thought were comfortable interacting with Blacks, their implicit anti-Black bias was negatively correlated with liking of those White targets—irrespective of whether those White targets displayed positive or negative nonverbal behavior

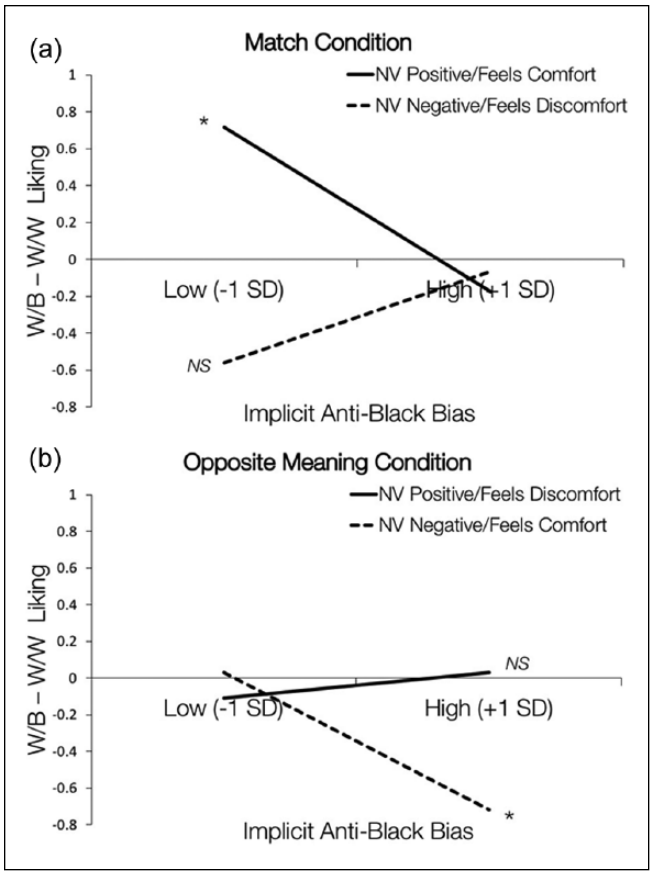

Figure 2. Experiment 2: Effect of implicit anti-Black bias on liking of White targets nonverbally comfortable or uncomfortable with a Black (W/B) versus a White (W/W) interaction partner, by instructions condition—(a) match instructions condition (nonverbal behaviors match feelings) and (b) opposite meaning instructions condition (nonverbal behaviors are opposite of feelings).
Note. Higher numbers indicate increasing preference for Whites with Blacks relative to Whites with Whites.

toward their Black partners. Also replicating Experiment 1, implicit bias did not predict liking when targets were thought to feel uncomfortable, irrespective of the tenor of their nonverbal behavior, though in these cases, the nonsignificant relationship was again positive. Therefore, participants' evaluations were influenced less by the objective treatment of Black interaction partners and more by what they thought ingroup targets' feeling were when interacting with Blacks.

Experiment 3

Experiments 1 and 2 consistently found support for the hypothesis that Whites' implicit anti-Black prejudice can predict liking of fellow Whites who seem comfortable with Blacks. According to the logic underlying our hypothesis, implicit bias will predict liking to the extent that participants infer correspondence between the ingroup targets they are rating and themselves. In Experiment 3, we measured White participants' perceived similarity to White targets comfortable with Blacks to test whether it mediated the relationship between their implicit anti-Black bias and liking of targets.

For this experiment, we also used a different measure of implicit bias. Because the ST-IAT used in Experiments 1 and 2 solely measures biases toward Blacks, it could be the case that the observed effect is better explained by implicit ingroup favoritism. The measure used in this study captured both anti-Black and pro-White implicit biases, allowing us to test whether the relationship persisted when controlling for implicit ingroup favoritism, in addition to explicit bias. Furthermore, the measure used in this study was subliminal, ensuring that participants were unaware we were assessing their racial bias.

Method

Participants. White undergraduates (n = 42; 48% male) were recruited for US$7 pay or course credit over one semester to complete this study in the lab. For such a sample to be powered at .8, the effect size, f, of implicit bias on liking of would need to be .44, suggesting our sample may have been underpowered.

Procedure. Participants first completed the subliminal priming measure of implicit bias. After this, participants completed a video ratings task similar to that used in Experiments 1 and 2, but with all comfortable videos. Finally, participants completed explicit measures and demographics.

Materials

Sequential subliminal priming measure. The subliminal priming task (Lowery, Hardin, & Sinclair, 2001) consisted of 128 trials in which participants were asked to fixate on a dot in the middle of the screen. Black and White pictures of 64 Black (32 female) and 64 White (32 female) faces were then presented parafoveally for 17 ms in random order at the four corners of the screen and then covered with a picture of a sunflower as a backward mask. Next, the word *good* or *bad* appeared at the center of the screen where the fixation dot had been. The word remained until participants identified which word had appeared by pressing a corresponding key on the keyboard (e.g., the *K* key for *good* and *D* key for *bad*). The word to which each key corresponded was counterbalanced across participants. Reaction time was recorded from the onset of the word *good* or *bad* until participants provided their response.

We also scored our task in an identical manner to Lowery and colleagues (2001). Mean reaction times to the words *good* and *bad* as a function of exposure to Black and White faces were recorded to serve as our measure of automatic racial bias. Reaction times less than 300 ms or greater than 3,000 ms were coded as missing values, and the remaining values were log-transformed. We then calculated one score

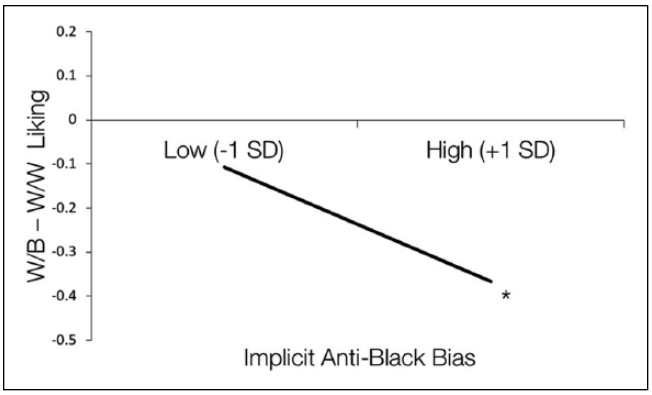

Figure 3. Experiment 3: Effect of implicit anti-Black bias on liking of ingroup targets with moderate nonverbal comfort toward a Black (W/B) versus a White (W/W) interaction partner.
Note. Higher numbers indicate increasing preference for Whites with Blacks relative to Whites with Whites.

for implicit anti-Black bias and a second score for implicit pro-White bias. Implicit anti-Black bias, in which higher numbers indicated greater negativity toward Blacks, was calculated by subtracting the response time for the word *bad* when primed with a Black face from response time to the word *good* when primed with a Black face. Implicit pro-White bias, in which higher numbers indicated greater positivity toward Whites, was calculated by subtracting response time to the word *good* when primed with a White face from response time to the word *bad* when primed with a White face. Implicit anti-Black and pro-White bias scores were negatively correlated ($r = -.30$, $p = .05$).

Interaction videos. We selected a different set of videos from the same previous study (Mallet et al., 2008). Selecting new videos enabled us to confirm that the findings were not idiosyncratic to the stimuli from Experiments 1 and 2. Each participant saw a total of eight silent, 90-s video clips of comfortable interactions, presented out of a pool of 12 such videos using a Latin Square design. Videos were divided equally by partner race (Black vs. White). Four coders (one male) rated targets' nonverbal comfort on three dimensions (i.e., comfort, friendliness, and rapport; 1 = *not at all comfortable* to 7 = *very comfortable*; $\alpha = .86$). All targets in the videos exhibited moderate nonverbal comfort with their partner regardless of the race of the partner ($M_{Black} = 4.2$, $SD_{Black} = 0.4$; $M_{White} = 4.2$, $SD_{White} = 0.7$).

Liking of target. Liking was assessed with four questions on a 6-point Likert-type scale (1 = *strongly disagree*; 6 = *strongly agree*). The four questions were "I like this person," "I feel warmly toward this person," "I would expect to get along easily with this person," and "I would expect to have a smooth interaction with this person" ($\alpha = .91$).

Perceived similarity to target. Perceived similarity was assessed with two questions ("I expect that I have a lot in common with this person" and "I expect this person and I share similar interests"; $\alpha = .94$) on a 6-point Likert-type scale (1 = *strongly disagree*; 6 = *strongly agree*).

Explicit prejudice measure. After the liking and perceived similarity ratings, participants completed the 20-item Ambivalent Racism Scale (Katz & Hass, 1988; 1 = *strongly disagree*, 8 = *strongly agree*), from which explicit anti-Black ($\alpha = .71$) and pro-Black ($\alpha = .87$) attitude scores were calculated.

Results

Mean-centered implicit anti-Black bias, implicit pro-White bias, and explicit anti-Black and explicit pro-Black prejudice scores were entered as predictors in the model. For the dependent variables, we calculated difference scores for both liking and perceived similarity, with ratings of targets with White partners subtracted from ratings of targets with Black partners. Higher scores indicated greater liking of and perceived similarity to targets interacting with Blacks relative to targets interacting with Whites.[5]

Liking. As predicted, participants' implicit anti-Black bias was negatively correlated with liking of White targets who seemed comfortable interacting with Blacks ($B = -2.32$, $SE = 1.02$, Wald $\chi^2 = 5.17$, $p = .023$, 95% CI = [−4.33, −0.32]; see Figure 3).

Perceived similarity. As predicted, participants' implicit anti-Black bias was also negatively correlated with feelings of similarity toward White targets who seemed comfortable interacting with Blacks ($B = -3.31$, $SE = 1.40$, Wald $\chi^2 = 5.57$, $p = .018$, 95% CI = [−6.07, −0.56]).

Mediation analyses. We conducted a mediation analysis to test whether perceived similarity could explain the effect of implicit anti-Black bias on liking of targets with Black partners compared with those with White partners. Using Preacher and Hayes's (2008) bootstrapping method, implicit anti-Black bias was entered as the predictor, the liking difference score was entered as the outcome (i.e., liking of targets with Black partners minus targets with White partners), the perceived similarity difference score was entered as the mediator (i.e., similarity to targets with Black partners minus targets with White partners), and implicit pro-White, explicit anti-Black, and explicit pro-White biases were entered as covariates. Using 10,000 bootstrapped samples, the unstandardized indirect effect of implicit anti-Black bias on liking via perceived similarity was significant ($B = -1.81$, $SE = 1.00$, 95% CI = [−4.30, −0.32]). Furthermore, the direct effect of implicit anti-Black bias on liking was reduced to nonsignificance ($B = -.51$, $SE = .97$, 95% CI = [−2.49, 1.46]; see Figure 4).[6]

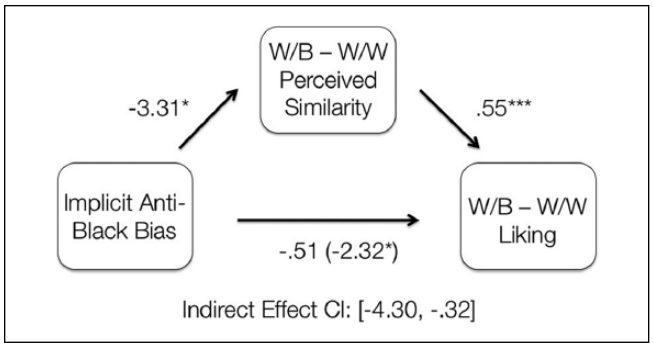

Figure 4. Experiment 3: Correlational model with perceived similarity to the target mediating the relationship between implicit anti-Black bias and liking of ingroup targets with moderate nonverbal comfort toward a Black interaction partner (W/B) minus ingroup targets with a White interaction partner (W/W).
Note. Coefficients are unstandardized.

Discussion

As hypothesized, we found evidence that White participants' implicit anti-Black bias predicts liking of ingroup targets to the extent that participants infer correspondence between targets and themselves. We found that participants' implicit anti-Black bias negatively correlated with both liking of and perceived similarity to targets comfortable with a Black interaction partner. Participants' ratings of perceived similarity with targets also mediated the relationship between their implicit bias and liking. Furthermore, we found that these effects persisted over and above implicit pro-White bias, suggesting that they are not driven by implicit ingroup favoritism. We also once again found that these effects persisted over and above the effects of consciously avowed prejudice.

General Discussion

In the present research, we turned from the well-studied experience of individuals within intergroup interactions (Dovidio, Hebl, Richeson, & Shelton, 2006) to examine the relatively understudied perceptions of individuals observing such interactions. We proposed the novel hypothesis that observers' liking of ingroup targets viewed in intergroup interactions is driven both by observers' implicit outgroup bias and targets' comfort with outgroup interaction. Specifically, we predicted that Whites' implicit anti-bias would negatively correlate with their liking of fellow Whites they see comfortably interacting with a Black stranger, but that this relationship should weaken or reverse when the ingroup members are observed to be uncomfortable with Blacks. Furthermore, this relationship was expected to hold over and above explicit racial bias, explicit motivation to be nonprejudiced, and implicit ingroup favoritism.

We found consistent support for this hypothesis across three experiments. As expected, Whites' implicit anti-Black bias was negatively correlated with their liking of ingroup members who were comfortable interacting with Blacks (Experiments 1-3). Across all experiments, this relationship held over and above conscious egalitarian attitudes and motivations as well as implicit ingroup favoritism. Also consistent with our expectations, this relationship was not significant when participants observed ingroup members who appeared uncomfortable with Blacks (Experiments 1-2). Although another prediction was that greater implicit bias could have predicted greater liking for uncomfortable targets, we identified two possible explanations for the null effect: First, the obviously uncomfortable targets were generally less liked by all participants, as is common (Alden & Wallace, 1995; Jones & Carpenter, 1986), and this general dislike may have overridden other interpersonal impressions. Second, expressing liking of a possibly prejudiced individual is usually seen as strongly counternormative (Crandall et al., 2002), possibly reducing participants' willingness to indicate affinity with targets who are thought to be uncomfortable with Blacks. Nonetheless, the questions of if and when higher implicit bias ever leads to greater liking of targets who are uncomfortable with outgroup contact merit further study.

Central to our hypothesis, we also found evidence that observers' implicit anti-Black bias predicted liking of ingroup targets seen interacting with Blacks to the extent that they inferred correspondence between the targets and themselves. Specifically, we found that observers' implicit anti-Black bias negatively correlated with liking of ingroup targets who were thought to be comfortable interacting with a Black partner, regardless of targets' nonverbal warmth or coldness toward their partner (Experiment 2). More directly, we found that perceived similarity to ingroup targets mediated the relationship between implicit anti-Black bias and liking of ingroup targets who were observed comfortably interacting with a Black person (Experiment 3).

In testing our hypothesis, we ruled out several viable alternative explanations of the findings. First, we ruled out the alternative stigma by association perspective, which suggests that observers will dislike ingroup targets' for merely being in proximity to outgroup members (e.g., Hebl & Mannix, 2003) irrespective of the observers' attitudes or the ingroup targets' apparent comfort. Second, we addressed the prediction that White observers' implicit bias might simply predict a preference for how Blacks should be treated. In this scenario, participants with lower implicit anti-Black bias would like targets who behaved friendlier toward Blacks while those with higher implicit bias would like targets who were more prejudicial toward Blacks. However, Experiment 2 directly demonstrated that observers' liking of ingroup targets depended on their perceptions of ingroup targets' felt comfort with interracial contact rather than their objective treatment of Black interaction partners. Finally, in Experiment 3, we entered both implicit anti-Black bias and implicit pro-White bias into our analysis, thereby ruling out a third alternative prediction that observers' liking of ingroup targets would be better explained by implicit ingroup favoritism.

Our perspective provides a novel framework for how intergroup interactions shape ingroup perceptions and has several theoretical and practical implications. First, the present work suggests that observers' inferences of targets' comfort may be a heretofore unidentified mechanism in past stigma by association research. The most commonly assumed mechanism in this past work is that negative perceptions of stigmatized outgroup members spontaneously transfer from the stigmatized person onto the nonstigmatized ingroup target (Hebl & Mannix, 2003; Walther, 2002). However, that research focuses on scenarios that imply ingroup targets are comfortable with the stigmatized outgroup member. For example, the majority of this literature tests stigma spread across family and friendship ties (e.g., Neuberg et al., 1994), which naturally imply familiarity and comfort. A smaller subset of research demonstrates that individuals are sometimes devalued for benignly being in the presence of a coincidental outgroup associate. For example, Pryor and colleagues (2012) found that Whites were rated more harshly for interacting with a Black stranger at a work function. The present findings hint that these past results might in part be explained by observers' assumptions of ingroup targets' comfort with the outgroup and whether targets' comfort corresponds with their own.

In addition, the present research extends beyond the extant stigma by association literature by showing that contact with stigmatized outgroup members can produce *either* negative or positive perceptions of an ingroup target. In Experiments 1 and 2, in addition to finding that participants' implicit bias negatively correlated with liking of targets comfortable with a Black partner, we also found instances wherein low bias participants demonstrated an overall preference for such targets (see Figures 1 and 2). Conversely, across all studies, we found instances wherein participants with high implicit bias had a clear aversion to Whites who were comfortable with Blacks (see Figures 1-3). These two outcomes are in line with past work showing that perceptions of similarity and dissimilarity both influence affiliation, but in opposite directions (Chen & Kenrick, 2002; Singh & Ho, 2000). More interestingly, the present studies suggest that both perceptions might be influential when viewing ingroup members in intergroup interactions. However, one limitation of the present work is that it did not examine how such perceptions might affect liking of the outgroup person seen interacting with the ingroup target. Future work would profit from testing the effect of participants' implicit bias on ratings of the outgroup member as a function of the ingroup interaction partner's perceived comfort with them.

The current work also highlights a mechanism by which homogeneous social networks can be formed and perpetuated. A growing body of research shows that network homogeneity is determined to a considerable extent by whether one has ingroup friends who act as conduits to and exemplars of positive interracial contact (Jacoby-Senghor et al., 2015; Zhou et al., 2018). For example, extended contact research highlights that one's likelihood of affiliating with outgroup members is increased by exposure to ingroup members who themselves have positive intergroup contact (Christ et al., 2010; Wimmer & Lewis, 2010; Wright, Aron, McLaughlin-Volpe, & Ropp, 1997), especially when opportunities for direct intergroup contact are rare (Eller, Abrams, & Gomez, 2012). Furthermore, past research has connected initial affiliative preferences, like those studied in the present work, with meaningful downstream affiliation and friendships (Montoya, Horton, & Kirchner, 2008; Neimeyer & Mitchell, 1988). Therefore, to the extent that people's implicit biases lead them to disproportionately befriend those who have similar comfort with intergroup contact, the relative homogeneity or heterogeneity of their social networks should persist or even intensify. Such a process could have far-reaching consequences, including in determining individuals' levels of intergroup bias (Van Laar, Levin, Sinclair, & Sidanius, 2005), effective problem-solving in work groups (Page, 2008; Phillips & Loyd, 2006), and the flow of social capital to marginalized individuals (DiTomaso, 2013; Lin, 1999). Tests of the current hypothesis in actual social networks would be valuable in determining how influential this mechanism is in network formation, development, and function.

In looking toward future research on the present topic, it is noteworthy that at times targets' feelings of comfort may diverge from their outward behavior in ways that lead observers to form incorrect impressions of targets' comfort. Although considerable research has shown that higher implicit bias relates to more negative, uncomfortable nonverbal behaviors (e.g., Bessenoff & Sherman, 2000; Dasgupta & Rivera, 2006; Dovidio et al., 2002; Dovidio, Kawakami, Johnson, Johnson, & Howard, 1997; Fazio, Jackson, Dunton, & Williams, 1995; McConnell & Leibold, 2001), the opposite may be true at times when evaluative concerns are high (e.g., Devine, Evett, & Vasquez-Suson, 1996; Shelton, Richeson, Salvatore, & Trawalter, 2005; Vorauer & Turpie, 2004). For example, Shelton et al. (2005) found that when discussing race with a Black partner, Whites higher, versus lower, in implicit bias were rated more favorably by their partner. Ironically, perceivers may therefore find themselves drawn to ingroup targets who have divergent levels of comfort with outgroup contact, particularly when targets' evaluative concerns are high. That said, this kind of self-regulation of one's behavior likely depends on surmounting several constraints, including having adequate cognitive resources to self-regulate (Mendes & Koslov, 2013), possessing an awareness of one's own biased behavior (Dovidio & Gaertner, 2004), and lacking any contextual justification for such behavior (Crandall & Eshleman, 2003; Monin & Miller, 2001). Future work should explore the extent to which evaluative concerns and deliberative expressions of prejudice or egalitarianism (e.g., verbal communication) influence the current findings.

In conclusion, the present research demonstrates that the social consequences of intergroup interactions extend far

beyond those directly involved. In fact, the tenor of these encounters appears to reverberate past the dyad, influencing the people who observe them in subtle and unconscious ways that may nonetheless have profound social implications.

Acknowledgment

Thanks to Andreana Kendrick, Nicole Shelton, Nick Turk-Browne, and Andrew Conway for their critical feedback; Allie Hootnick and Kelly Hoffman for their early contributions; and the numerous research assistants who aided in data collection and manuscript preparation.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. Experiment 1 was the second study conducted. In response to reviewer notes and for the purposes of narrative clarity, the chronologically first study has been moved to Experiment 3.
2. For the sake of clarity, only the statistics relevant to our hypothesis are reported in the article. However, full effect statistics, including those for forms of bias other than implicit anti-Black bias, are provided in supplementary materials. In general, the inclusion or exclusion of alternative forms of bias did not change the interpretation of our findings.
3. See descriptive statistics in Supplemental Table S1. See Supplemental Table S2 for full reporting of effects of explicit prejudice and IMS. Results excluding IMS can also be found in the aforementioned table. In Experiments 2 and 3, the overall patterns of results do not change as a function of the inclusion of IMS.
4. See descriptive statistics in Supplemental Table S4. See Supplemental Table S5 for full reporting of effects of explicit prejudice and IMS.
5. See descriptive statistics in Supplemental Table S7. See Supplemental Table S8 for full reporting of effects of explicit prejudice and implicit ingroup favoritism.
6. See Supplemental Table S9 for full reporting of effects of explicit prejudice and implicit ingroup favoritism in the mediation analysis.

Supplemental Material

Supplemental material is available online with this article.

References

Alden, L. E., & Wallace, S. T. (1995). Social phobia and social appraisal in successful and unsuccessful social interactions. *Behaviour Research and Therapy*, *33*, 497-505.

Berscheid, E. (1985). Interpersonal attraction. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (3rd ed., pp. 413-484). New York, NY: Random House.

Bessenoff, G. R., & Sherman, J. W. (2000). Automatic and controlled components of prejudice toward fat people: Evaluation versus stereotype activation. *Social Cognition*, *18*, 329-353.

Blascovich, J., Mendes, W. B., Hunter, S. B., Lickel, B., & Kowai-Bell, N. (2001). Perceiver threat in social interactions with stigmatized others. *Journal of Personality and Social Psychology*, *80*, 253-267.

Bluemke, M., & Friese, M. (2008). Reliability and validity of the Single-Target IAT (ST-IAT): Assessing automatic affect towards multiple attitude objects. *European Journal of Social Psychology*, *38*, 977-997.

Castelli, L., De Amicis, L., & Sherman, S. J. (2007). The loyal member effect: On the preference for ingroup members who engage in exclusive relations with the ingroup. *Developmental Psychology*, *43*, 1347-1359.

Chen, F. F., & Kenrick, D. T. (2002). Repulsion or attraction? Group membership and assumed attitude similarity. *Journal of Personality and Social Psychology*, *83*, 111-125.

Christ, O., Hewstone, M., Tausch, N., Wagner, U., Voci, A., Hughes, J., & Cairns, E. (2010). Direct contact as a moderator of extended contact effects: Cross-sectional and longitudinal impact on outgroup attitudes, behavioral intentions, and attitude certainty. *Personality and Social Psychology Bulletin*, *36*, 1662-1674.

Conley, T. D., Rabinowitz, J. L., & Hardin, C. D. (2010). O. J. Simpson as shared (and unshared) reality: The impact of consensually shared beliefs on interpersonal perceptions and task performance in different- and same-ethnicity dyads. *Journal of Personality and Social Psychology*, *99*, 452-466.

Crandall, C. S., & Eshleman, A. (2003). A justification-suppression model of the expression and experience of prejudice. *Psychological Bulletin*, *129*, 414-446.

Crandall, C. S., Eshleman, A., & O'Brien, L. (2002). Social norms and the expression and suppression of prejudice: The struggle for internalization. *Journal of Personality and Social Psychology*, *82*, 359-378.

Dasgupta, N. (2004). Implicit ingroup favoritism, outgroup favoritism, and their behavioral manifestations. *Social Justice Research*, *17*(2), 143-169.

Dasgupta, N., & Rivera, L. M. (2006). From automatic anti-gay prejudice to behavior: The moderating role of conscious beliefs about gender and behavioral control. *Journal of Personality and Social Psychology*, *91*, 268-280.

Devine, P. G., Evett, S. R., & Vasquez-Suson, K. A. (1996). Exploring the interpersonal dynamics of intergroup contact. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition. Vol. 3: The interpersonal context* (pp. 423-464). New York, NY: Guilford Press.

DiTomaso, N. (2013). *The American non-dilemma: Racial inequality without racism*. New York, NY: Russell Sage Foundation.

Dixon, J., Tredoux, C., Durrheim, K., Finchilescu, G., & Clack, B. (2008). The inner citadels of the color line: Mapping the microecology of racial segregation in everyday life spaces. *Social and Personality Psychology Compass*, *2*, 1547-1569.

Dovidio, J. F., & Gaertner, S. L. (2004). Aversive racism. *Advances in Experimental Social Psychology*, *36*, 1-52.

Dovidio, J. F., Hebl, M., Richeson, J. A., & Shelton, J. N. (2006). Nonverbal communication, race, and intergroup interaction. In V. Manusov & M. L. Patterson (Eds.), *The Sage handbook of nonverbal communication*. Thousand Oaks, CA: Sage (pp. 481-500).

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of Personality and Social Psychology*, *84*, 754-770.

Dovidio, J. F., Kawakami, K., Johnson, C., Johnson, B., & Howard, A. (1997). On the nature of prejudice: Automatic and controlled processes. *Journal of Experimental Social Psychology*, *33*, 510-540.

Eller, A., Abrams, D., & Gomez, A. (2012). When the direct route is blocked: The extended contact pathway to improving intergroup relations. *International Journal of Intercultural Relations*, *5*, 637-646.

Fazio, R. H., Jackson, J. R., Dunton, B. C., & Williams, C. J. (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes. A bona fide pipeline? *Journal of Personality and Social Psychology*, *69*, 1013-1027.

Goffman, E. (1963). *Stigma: Notes on the management of spoiled identity*. Englewood Cliffs, NJ: Prentice-Hall.

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and Using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology*, *97*, 17-41.

Hahn, A., Judd, C. M., Hirsh, H. K., & Blair, I. V. (2014). Awareness of implicit attitudes. *Journal of Experimental Psychology: General*, *143*, 1369-1392.

Hebl, M. R., & Mannix, L. M. (2003). The weight of obesity in evaluating others: A mere proximity effect. *Personality and Social Psychology Bulletin*, *29*, 28-38.

Jacoby-Senghor, D. S., Sinclair, S., & Smith, C. T. (2015). When bias binds: Effect of implicit outgroup bias on ingroup affiliation. *Journal of Personality and Social Psychology*, *109*, 415-433.

Jones, W. H., & Carpenter, B. N. (1986). Shyness, social behavior, and relationships. In W. H. Jones, J. M. Cheek, & S. R. Briggs (Eds.), *Shyness: Perspectives on research and treatment* (pp. 227-238). New York, NY: Plenum Press.

Judd, C. M., Kenny, D. A., & McClelland, G. H. (2001). Estimating and testing mediation and moderation in within-subject designs. *Psychological Methods*, *6*, 115-134.

Katz, I., & Hass, R. G. (1988). Racial ambivalence and American value conflict: Correlational and priming studies of dual cognitive structures. *Journal of Personality and Social Psychology*, *55*, 893-905.

Lin, N. (1999). Building a network theory of social capital. *Connections*, *22*(1), 28-51.

Lowery, B. S., Hardin, C. D., & Sinclair, S. (2001). Social influence effects on automatic racial prejudice. *Journal of Personality and Social Psychology*, *5*, 842-855.

Mallet, R. K., Wilson, T. D., & Gilbert, D. T. (2008). Expect the unexpected: Failure to anticipate similarities leads to an intergroup forecasting error. *Journal of Personality and Social Psychology*, *94*, 265-277.

McConnell, A. R., & Leibold, J. M. (2001). Relations among the Implicit Association Test, explicit attitudes, and discriminatory behavior. *Journal of Experimental Social Psychology*, *37*, 435-442.

Mendes, W. B., & Koslov, K. (2013). Brittle smiles: Positive biases toward stigmatized and outgroup targets. *Journal of Experimental Psychology: General*, *142*, 923-933.

Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology*, *81*, 33-43.

Montoya, R. M., Horton, R. S., & Kirchner, J. (2008). Is actual similarity necessary for attraction? A meta-analysis of actual and perceived similarity. *Journal of Social and Personal Relationships*, *25*, 889-922.

Neimeyer, R. A., & Mitchell, K. A. (1988). Similarity and attraction: A longitudinal study. *Journal of Social and Personal Relationships*, *5*, 131-148.

Neuberg, S. L., Smith, D. M., Hoffman, J. C., & Russell, F. J. (1994). When we observe stigmatized and "normal" individuals interacting: Stigma-by-association. *Personality and Social Psychology Bulletin*, *20*, 196-209.

North, M. S., & Fiske, S. T. (2013). Subtyping ageism: Policy issues in succession and consumption. *Social Issues and Policy Review*, *7*, 36-57.

Nosek, B. A., & Smyth, F. L. (2007). A multitrait-multimethod validation of the Implicit Association Test: Implicit and explicit attitudes are related but distinct constructs. *Experimental Psychology*, *54*, 14-29.

Page, S. E. (2008). *The difference: How the power of diversity creates better groups, firms, schools, and societies*. Princeton, NJ: Princeton University Press.

Phillips, K. W., & Loyd, D. L. (2006). When surface and deep-level diversity collide: The effects on dissenting group members. *Organizational Behavior and Human Decision Processes*, *99*, 143-160.

Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology*, *75*, 811-848.

Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, *40*, 879-891.

Pryor, J. B., Reeder, G. D., & Monroe, A. E. (2012). The infection of bad company: Stigma-by-association. *Journal of Personality and Social Psychology*, *102*, 224-241.

Ranganath, K. A., Smith, C. T., & Nosek, B. A. (2008). Distinguishing automatic and controlled components of attitudes from direct and indirect measurement methods. *Journal of Experimental Social Psychology*, *44*, 382-396.

Richeson, J. A., & Shelton, J. N. (2003). When prejudice does not pay: Effects of interracial contact on executive function. *Psychological Science*, *14*, 287-290.

Richeson, J. A., & Shelton, J. N. (2005). Brief report: Thin slices of racial bias. *Journal of Nonverbal Behavior*, *29*, 75-86.

Rozin, P., Markwith, M., & Nemeroff, C. (1992). Magical contagion beliefs and fear of AIDS. *Journal of Applied Social Psychology*, *22*, 1081-1092.

Seery, M. D., Holman, E. A., & Silver, R. C. (2010). Whatever does not kill us: Cumulative lifetime adversity, vulnerability, and resilience. *Journal of Personality and Social Psychology*, *99*, 1025-1041.

Shelton, J. N., Richeson, J. A., Salvatore, J., & Trawalter, S. (2005). Ironic effects of racial bias during interracial interactions. *Psychological Science*, *16*, 397-402.

Sigelman, C. K., Howell, J. L., Cornell, D. P., Cutright, J. D., & Dewey, J. C. (1991). Courtesy stigma: The social implications of associating with a gay person. *The Journal of Social Psychology*, *131*, 45-56.

Singh, R., & Ho, S. Y. (2000). Attitudes and attraction: A new test of the attraction, repulsion and similarity-dissimilarity and

asymmetry hypotheses. *British Journal of Social Psychology*, *39*, 197-211.

Towles-Schwen, T., & Fazio, R. H. (2003). Choosing social situations: The relation between automatically activated racial attitudes and anticipated comfort interacting with African Americans. *Personality and Social Psychology Bulletin*, *29*, 170-182.

Van Laar, C., Levin, S., Sinclair, S., & Sidanius, J. (2005). The effect of university roommate contact on ethnic attitudes and behavior. *Journal of Experimental Social Psychology*, *41*, 329-345.

Vorauer, J. D., & Turpie, C. A. (2004). Disruptive effects of vigilance on dominant group members' treatment of outgroup members: Choking versus shining under pressure. *Journal of Personality and Social Psychology*, *87*, 384-399.

Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology*, *82*, 919-934.

Wimmer, A., & Lewis, K. (2010). Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. *American Journal of Sociology*, *116*, 583-642.

Word, C. O., Zanna, M. P., & Cooper, J. (1974). The nonverbal mediation of self-fulfilling prophecies in interracial interaction. *Journal of Experimental Social Psychology*, *10*, 109-120.

Wright, S. C., Aron, A., McLaughlin-Volpe, T., & Ropp, S. A. (1997). The extended contact effect: Knowledge of cross-group friendships and prejudice. *Journal of Personality and Social Psychology*, *73*, 73-90.

Zeger, S. L., & Liang, K. Y. (1986). Longitudinal data analysis for discrete and continuous outcomes. *Biometrics*, *42*, 121-130.

Zhou, S., Page-Gould, E., Aron, A., Moyer, A., & Hewstone, M. (2018). The Extended Contact Hypothesis: A Meta-Analysis on 20 Years of Research. *Personality and Social Psychology Review*.